

The EM algorithm and its extensions.

Geoffrey J. MCLACHLAN and Thriyambakam KRISHNAN. New York: Wiley, 1996. xvii+274 pp. \$59.95.

MacLachlan and Krishnan remark in their foreword that "It is surprising that ... no book (on EM) has appeared". This is perhaps not too remarkable due to the fact that the seminal paper by Dempster, Laird and Rubin served the purpose of synthesizing and energizing applications and research in the area and so no book length treatment was needed for many years. However, given the recent pace of development of EM theory, applications and extensions, the appearance of this well-written, book-length treatment is now appropriate.

Who might this book benefit? The book is clearly intended as expository (it contains no exercises) and the authors suggest that the book will be useful to both practitioners and researchers of statistics and that the book will be accessible to graduates in Statistics (bachelors level??). This is a perhaps an overly optimistic reach. The book *does* contain numerous examples that lead the way in showing how the EM algorithm is to be applied. However, much of the book remains out of the reach of anyone except those with a bit of familiarity with the research literature.

In my opinion, the strongest contribution the book makes is in relating the various modifications and improvements of EM and in indicating relationships to alternate techniques. These are especially hard to establish on one's own since it would require gathering together research in different journals, by different authors, in different notations. I especially enjoyed the chapters on Standard Errors and Speeding Up Convergence (Chapter 4) and Extensions of the EM Algorithm (Chapter 5). The references are quite up to date.

I would have liked to have seen a bit more advice for the practitioner: general strategies for applying and implementing EM, stopping strategies, advice as to when EM can be expected to be better or worse than competing techniques, and some practical advice on when it can be expected to find a global maximum and/or how to help it do so. Some techniques seem to have been enthusiastically endorsed, even though the practitioner may have trouble implementing them. For example, for many problems, the Meng and Rubin (1991) method of obtaining standard errors can be numerically unstable.

Being slightly critical there were a few portions of the book I didn't like. The bootstrap section didn't really explain the bootstrap and didn't show how the bootstrap can be used to special advantage in missing data problems. The section on Newton-type methods claims that EM is "a method for finding zeros of a function" (it is really more a hill-climbing technique and often works to find the MLE even when it is not a zero of the likelihood function). Self-consistency was discussed and "proven" for the MLE without ever being formally defined.

That said, I would recommend the book to researchers wanting to learn more about the EM algorithm and variants. Practitioners and more applications oriented

statisticians may want to peruse the examples but may prefer to check it out of the library rather than own it.

Charles E. MCCULLOCH
Cornell University

REFERENCE

Meng, X.L., and Rubin, D.B. (1991). "Using EM to Obtain Asymptotic Variance-Covariance Matrices for Missing-Data Problems using EM." *Journal of the American Statistical Association*, 86, 899-909.